



# **Translation Memory Technology Assessment**

**Carol Van Ess-Dykema**

**Susan Converse**

**John S. White**

National Virtual Translation Center

**Dennis Perzanowski**

**Rachael Richardson**

**Tucker Maney**

Naval Research Laboratory

**MT Summit XII**

**Government User Session**

26-30 August, 2009

Ottawa, Ontario, Canada



Government User Session



## **Topics**

- **Introduction**
- **NVTC Requirements for Multi-Genre Translation**
- **Motivations for TM Technology Assessment**
- **Pilot Study**
- **Results**
- **Collaboration with NIST**



Government User Session

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE <b>AUG 2009</b>	2. REPORT TYPE		3. DATES COVERED <b>00-00-2009 to 00-00-2009</b>		
4. TITLE AND SUBTITLE <b>Translation Memory Technology Assessment</b>			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S)			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Naval Research Laboratory, Washington, DC, 20375</b>			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT <b>Same as Report (SAR)</b>	18. NUMBER OF PAGES <b>10</b>	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			



## NVTC Mission

NATIONAL  
VIRTUAL  
TRANSLATION  
CENTER

To provide timely, accurate translation services  
for the security of the nation



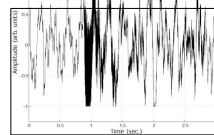
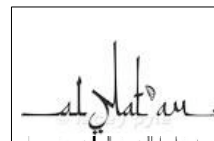
Government User Session



## NVTC Services

NATIONAL  
VIRTUAL  
TRANSLATION  
CENTER

- **Translation**
  - English to target language(s)
  - Foreign language(s) to target language(s) – usually English
- **Transcription**
  - Audio
  - Video
- **Summarization**
- **Gisting**



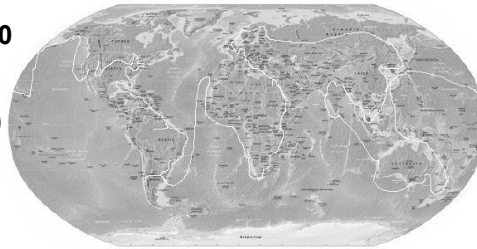
Government User Session



## NVTC Languages (2008)



- Translated for approx. 30 U.S. Government entities
- Translated from more than 50 Languages
- Translated into more than 70 languages



Government User Session



## Genres



### Examples:

- Book chapters
- Journal articles
- Web news articles
- Email messages
- PowerPoint slides
- Brochures
- TV broadcasts
- Radio broadcasts
- Maps



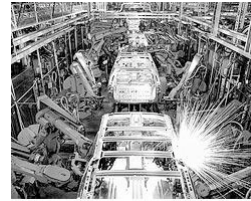
Government User Session



## Motivations for TM Technology Assessment

NATIONAL  
VIRTUAL  
TRANSLATION  
CENTER

- **Application of TM technology**
  - **Industry**
    - Automobile industry has documented return on investment for updating the translation of auto manufacturing manuals for off-shore assembly plants
    - Standardized terminology for Parts and Components Suppliers
  - **Example**
    - General Motors
    - Languages: English to German, European Spanish, Canadian French, Italian, Mexican Spanish and Dutch



Government User Session



## Motivations for TM Technology Assessment

NATIONAL  
VIRTUAL  
TRANSLATION  
CENTER

- **NVTC also translates manuals:**
  - U.S. Army field manuals from English to Arabic and other languages to train Iraqi Security Forces and other coalition partners.
- **NVTC translates many other genres as well:**
  - Journal articles, news stories, broadcast interviews, memoranda, emails, web sites, PowerPoint slides, etc.



Government User Session



## Translation Memory Technology Assessment: Pilot Study



### Goals:

- Develop TM assessment methodology using 3 common languages
- Develop candidate performance metrics for assessing TM usage across genres (e.g., translation speed, number of QC edits)
- Develop candidate linguistic parameters to characterize genres (e.g., sentence structure, morphology)
- Initiate development of TM memory banks for 3 languages

### End goal:

- To determine if and under what circumstances using TM technology will result in improvements in translation speed and accuracy when translating IC and DoD documents in the NVTC workflow



Government User Session



## NVTC – Naval Research Laboratory Joint Research Partnership to Assess TM Technology



- NVTC
  - Data for translation
  - Translators, quality control professionals
  - Researchers
- NRL
  - Data for translation
  - Researchers
  - Joint Research Agreement (in progress)
    - First NRL joint agreement with the IC
    - First NVTC joint agreement with US Government research entity
- NIST
  - Assessment of experimental methods
  - IC/DoD Workshops
  - Metrics development

**NIST**



Government User Session



## Translation Memory Technology Assessment: Pilot Study



Languages	Documents	Memory Bank Source Data
Arabic and Chinese	Transcripts of broadcast news	Transcripts of broadcast news
Russian	Military journal articles	Unknown source

For each language:

- One control document translated by hand
- Three test documents translated using TM tool
- Two translators each translate four documents (control and test)
- Two QC professionals check the four translations, one QC professional per translator



Government User Session



## Translation Memory Technology Assessment: Pilot Study



### Analysis Compare

- The quality of the translation for Arabic and Chinese.
  - Will highlight what linguistic criteria (e.g., sentence structure, morphology) might be relevant in characterizing genres.

### Compare

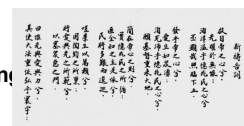
- The quality of the Russian translations to the Arabic and Chinese translations.
  - Chinese and Arabic will show the effect of using a memory bank built from documents of the same genre as the translation source documents
  - Russian will show the effect of using a memory bank that is populated with unrelated genre segments.

### Compare

- The control document translated “by hand” to the translation of the matching test document (comparable genre, length, and topic) using the TM tool.
  - Will give some measure of how much TM technology helps improve translation speed and accuracy for each language and genre combination.



Government User Session

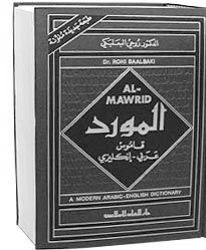




## What did we Measure?

NATIONAL  
VIRTUAL  
TRANSLATION  
CENTER

- **Translator:**
    - **Usability (quantitative)**
      - Mouse clicks
      - Tool use log
      - Memory bank updates
    - **User assessment (qualitative)**
      - Questionnaires
        - After manual and TM-assisted tasks
    - **Controls:**
      - Used single (hardcopy) dictionary per language
      - Trained on TM system
  - **Quality Control Professional:**
    - **User assessment (qualitative)**
      - Questionnaires on each translation QC'ed
- Controls: not told which translations were purely manual



Government User Session



## Questionnaire System

NATIONAL  
VIRTUAL  
TRANSLATION  
CENTER

Questionnaires were administered through a locally served web application.





## Questionnaire System



Apart from free-response prompts, all questionnaire items required the user to

Search Tool Evaluation : Question 1 of 7

http://localhost:8081/tes/eval/1

Search Tool Evaluation : Quest...

Logged in as User #7 for session 1.

**Session Evaluation: 1 of 7**

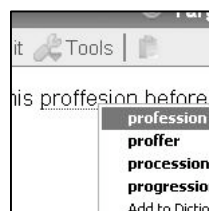
I had the resources I needed to produce a translation with proper grammar.

☐ No Response   ☐ Strongly Disagree   ☐ Disagree   ☐ Somewhat Disagree   ☐ Neither Agree nor Disagree   ☐ Somewhat Agree   ☐ Agree   ☐ Strongly Agree

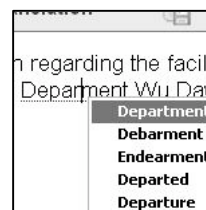
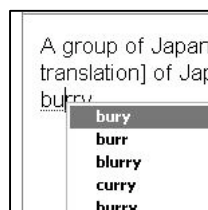
Done



Translators made extremely heavy use of the automatic spell-checker.  
We did not anticipate this behavior.



Russian-English translation.



Chinese-English translation.  
Government User Session





## Software Evaluation Questionnaire: Conclusions

NATIONAL  
VIRTUAL  
TRANSLATION  
CENTER

- Ignoring differences between sessions, users seemed to have a slightly favorable opinion of the TM system's:
  - Utility
  - ease of use
  - Information retrieval performance.
- This data is anecdotal, but suggests a trend:
  - Subjects' opinion of the system generally improved in later sessions
  - Further study with a larger group of subjects may produce statistically significant results.



Government User Session



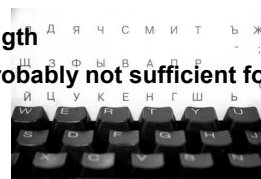
## Observations

NATIONAL  
VIRTUAL  
TRANSLATION  
CENTER

- Translators are used to multiple internet sources, colleagues, specialized glossaries for terminology support
  - Limiting terminology source intended as control, but creates potential bias
  - Lesson learned: allow more access to open sources
- Translators improved use of TM quickly over a few days
  - Lesson learned: longer ramp-up time
- Native memory bank (the Russian memory bank resident in the commercial TM) was insufficient for Russian military journal articles
  - Rarely scored matches of significant length
  - Lesson learned: native memory bank probably not sufficient for domain and/or genre specific text.



Government User Session





## Collaboration with the National Institute of Standards and Technology (NIST)



- NVTC-NRL will present the pilot study TM assessment methodology and results to NIST.
- NIST will validate and refine the outcomes from the pilot study, resulting in:
  - A set of candidate performance metrics for measuring the translation improvement when using TM technology;
  - A series of workshops for researchers, developers, users, and other TM stakeholders;
  - An augmentation of the MetricsMATR series, or similar forum for TM metrics
- Meanwhile, NVTC-NRL will begin a memory bank for a 4<sup>th</sup> language (possibly from Africa) to use in follow-up assessment experiments.



Government User Session



## Translation Memory Technology Assessment: Follow-On Study



### Follow-on Study

- Continue development of memory banks for all 4 languages
  - Arabic, Chinese, Russian, African language
- Add new genres
- Increase independence of experimental components
  - Pilot study measures effects of TM technology for assisting translators as individuals.
  - Follow on study will better distinguish between the translator and the TM technology as the direct source of improvement
- Study the structure and use of memory banks
  - Investigate organizing memory banks by genre, topic, linguistic structure, or other criteria

እሺ፡ እዚህ፡ ጠባቂዎች፡ አቡን፡ ሄጃ፡ ዳግመኛ፡  
እንደዚህ፡ እንዳይ፡ ደግሞ፡ አስረዳው፡ እለሁ፡  
ትንሽ፡ ሙድ፡ ታች፡ ከተገዘ፡ በጋላ፡ እየተራመደ፡  
ሲሄድ፡ አለሁ፡  
ደጋቢ፡ ጥ፡ ከአምንም፡ ፍርድታ፡ በእንባሰው፡ ጀርባ፡  
ሳይ፡ እረፈ፡ ግን፡ ፈቅዶልህ፡ ነው፡ ቢቁ፡ ሙተህ፡  
ልጆቻችን፡ የምታስፈራራው፡ እያለክጀርባው፡ ሳይ፡  
ሆኖ፡ በየደራ፡ ሆኖ፡  
እንበሰው፡ ግን፡ ከመራመዱ፡ አስቆሞውን፡



Government User Session